

Review

Statistical Binarization Techniques for Document Image Analysis

Saad M. Ismail, Siti Norul Huda Sheikh Abdullah, Fariza Fauzi

Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia

Article history

Received: 03-09-2017

Revised: 05-12-2017

Accepted: 03-01-2018

Corresponding Author:

Fariza Fauzi

Faculty of Information Science
and Technology, Universiti
Kebangsaan Malaysia,
Malaysia

Email: fariza.fauzi@ukm.edu.my

Abstract: Binarization is an important process in image enhancement and analysis. Currently, numerous binarization techniques have been reported in the literature. These binarization methods produce binary images from color or gray-level images. This article highlights an extensive review on various binarization approaches which are also referred to as thresholding methods. These methods are grouped into seven categories according to the employed features and techniques: histogram shape-based, clustering-based, entropy-based, object-attribute-based, spatial, local and hybrid methods. Most active binarization researchers exploit several initial information from the source image such as histogram shape, measurement space clustering, entropy, object attributes, spatial correlation and local gray level surface with a special attention to statistical information description features of image used in recent thresholding techniques.

Keywords: Binarization, Thresholding, Statistical Features, Document Image

Introduction

Binary image representation, a black and white representation of object and non-object (i.e. background) is the preferred format for image analysis especially document image analysis which consists of textual objects. Hence, in the application of document image analysis, binarization which converts color or gray-level images into binary images is a crucial pre-processing step. The accuracy of this binarization process affects the performance of the subsequent stages in detecting and extracting text from a document image which is the main purpose of performing binarization (or thresholding, the terms are used interchangeably throughout the article) on document images. Other purposes include noise removal and size reduction of the images in memory. Ultimately, this process increases the visibility of the important information in an image by discarding unimportant information (Kefali *et al.*, 2010). This preservation of important image information is achieved by reducing all levels in a color or gray-level image into two levels: black text and white background. Binarization usually employs manual or automatic threshold values. Thresholding outputs binary images by turning all gray-level pixels into either black pixels or white depending on some threshold value. Thresholding only considers pixel intensity and disregards relationships between the pixels which is a major limitation (Morse, 1998) whereby noise pixels can be mistakenly included and informative pixels but isolated from other pixels (e.g.

pixels near the boundaries of text region) can be missed (wrongly excluded). The effects are compounded as the noise increases because the true intensity of the flawed region can no longer be represented by pixel intensity.

Hence, current researches in thresholding methods are taking into account pixel relationship. There are three general categories of thresholding techniques: global, local and hybrid:

- 1) Global Thresholding: This threshold type, T depends only on a single threshold value to decide the change of gray-level pixels into black or white pixels
- 2) Local Thresholding: Adaptive thresholding divides an image into a number of regions. Each region performs a thresholding process based on a threshold value that is computed depending on the specific region contents. In other words, the threshold value adapts according to its local pixel information (within the region)
- 3) Hybrid Thresholding: Hybrid methods combine global and local information to label a pixel as object or background

Both local and global thresholding methods were long established to solve automatic binarization issues. However, in addressing the shortcomings of each method, research advances to hybrid binarization methods to cater to more challenging cases (Bataneh *et al.*, 2011; 2015; Kavallieratou and Stathis, 2006). Sezgin and Sankur (2004) refine the categorization according to the

different approaches and features used and introduce six classes of binarization methods: Histogram Shape-Based Methods, Clustering-Based Methods, Entropy-Based Methods, Object Attribute-Based Methods, the Spatial Methods and Local Methods:

- 1) Histogram shaped-based methods analyze properties of the smoothed histograms (i.e. peaks, valleys, curvatures)
- 2) Clustering-based methods cluster the gray-level samples into two parts – background and foreground, or alternatively, model the samples as a mixture of two Gaussians
- 3) Entropy-based methods utilize the entropy of the background and foreground regions, the cross-entropy between the original and binarized images, etc
- 4) Object attributed-based methods measure the similarity between the gray-level and the binarized images e.g. fuzzy shape similarity, edge coincidence, etc
- 5) Spatial methods use higher-order probability distribution and/or correlation between pixels
- 6) Local methods adapt the threshold value on each pixel according to the local image characteristics

Fig. 1 places the two categorizations of thresholding methods into perspective: the class of hybrid methods is positioned on the leftmost block corresponding to its definition that it is a class of methods that combines both global and local methods; Sezgin and Sankur (2004) categories are on the rightmost of the block diagram where the first four categories could be considered as global methods and the last two as local methods. Any combinations of Sezgin and Sankur (2004)'s categories could be considered as hybrid methods as well.

For document images, especially those of varying degraded quality and noise as shown in Fig. 2-4, binarization is a challenging task (Gatos *et al.*, 2008; Mahajan and Patil, 2015). Sample degraded text images are from DIBCO 2009 and 2014 datasets (Gatos *et al.*, 2009; Ntirogiannis *et al.*, 2014) have been compiled in Fig. 2 and magnified in Fig. 3 to illustrate the different degradations.

Degradations in document images are common. Examples of degradations and the possible causes include non-uniform background intensity (c.f. Fig. 4), low contrast, poor illumination, shadows, dark spots etc. caused by poor lighting; smears/stains, dirty, dark streaks/regions caused by ink bleeding, careless handling and storage etc. (c.f. Fig. 2 and 3).

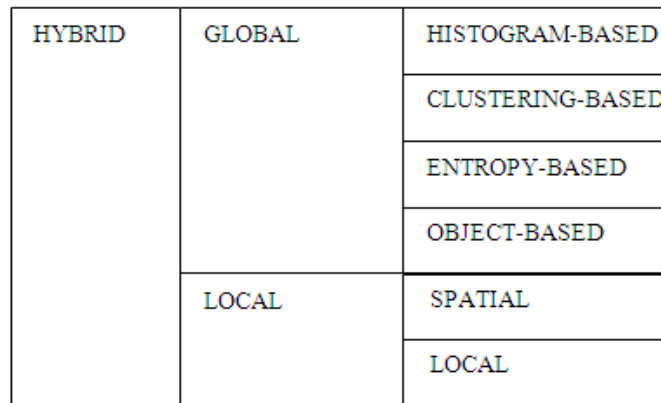
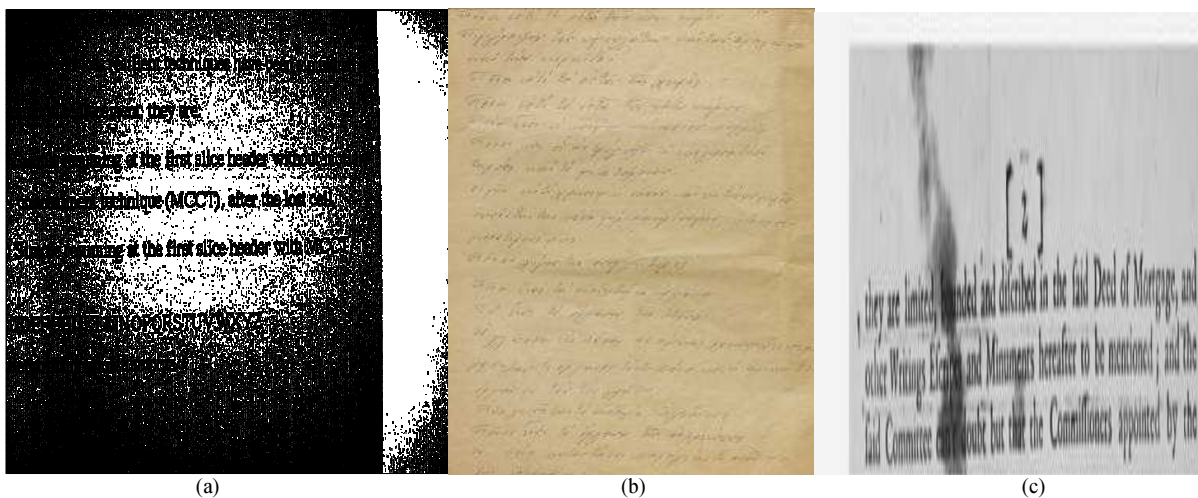


Fig. 1: Block diagram of the categories of binarization techniques



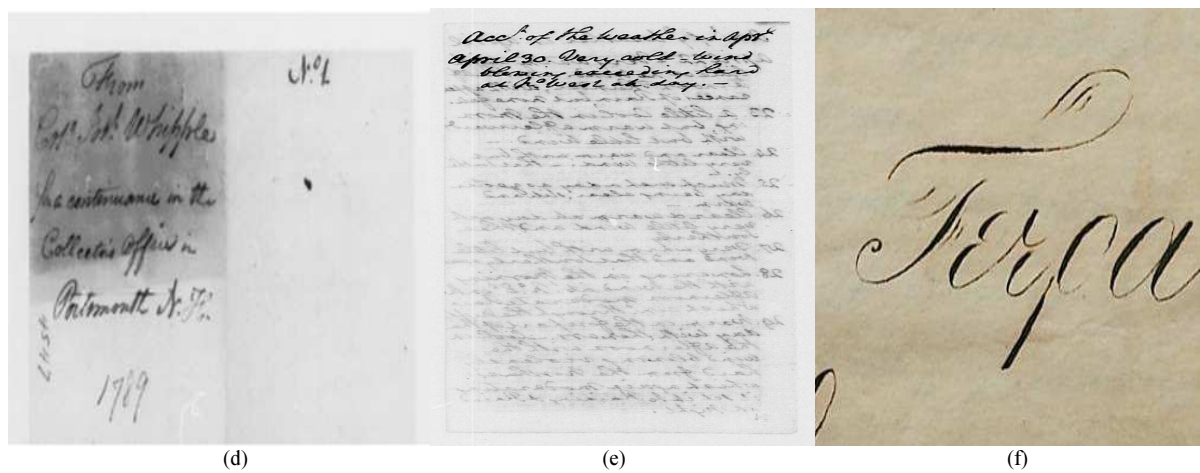


Fig. 2: Examples of degraded document images: (a) bad illumination, noise and shadow (b) a poor low-contrast image (c) dirty and dark spots (d) a non-uniform illumination with thin pen stroke (e) ink bleeding-through and (f) broken characters, light. Source: Gatos *et al.* (2009) and Ntirogiannis *et al.* (2014)

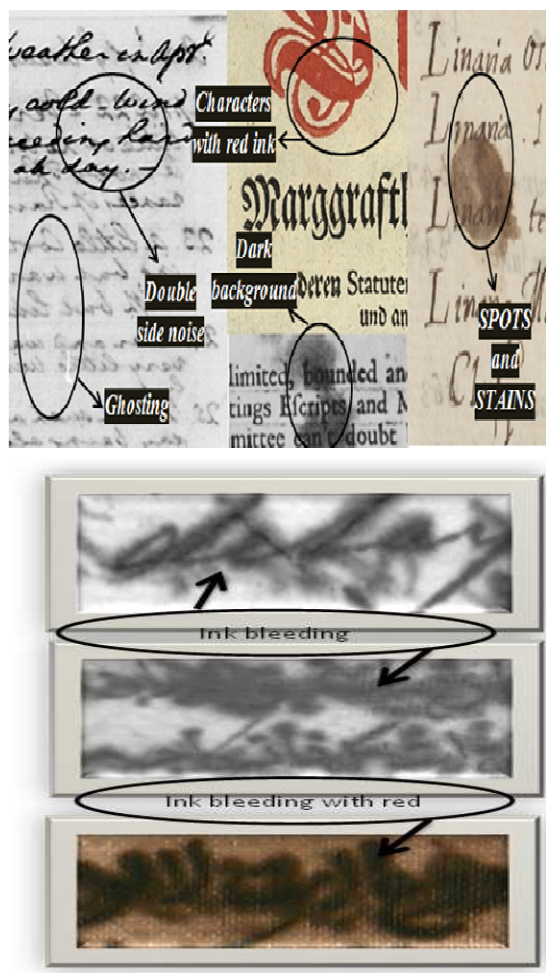


Fig. 3: Examples of degraded document images containing noises such as double side, ghosting, colour text, spots, stains and ink bleed. Source: Gatos *et al.* (2009) and Ntirogiannis *et al.* (2014)

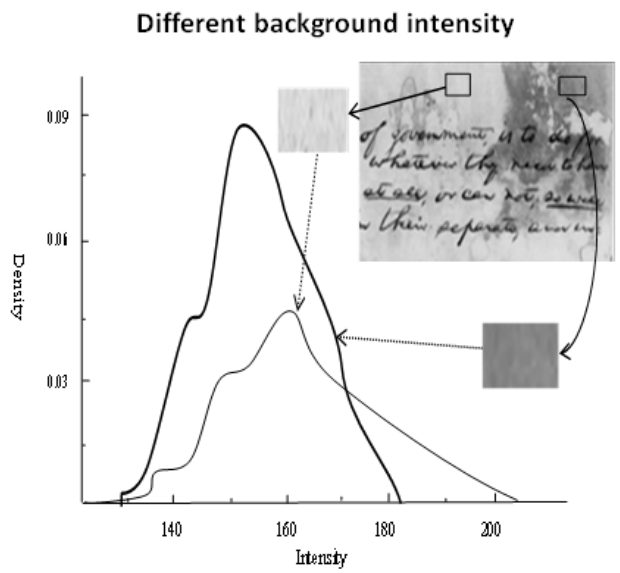


Fig. 4: An example of non-uniform background with different background intensity. Sample text image is from DIBCO 2009 dataset. Source: Gatos *et al.* (2009)

Some enhancement methods for degraded document image use its visual contents in order to improve the quality of the image, such as shape and texture information as well as some statistical information.

The objective of this article is to compare different methods of binarization involving document text images in relation to its condition involving pixel behavior. This article is organized into eight sections comprising of the introduction, an overview of thresholding methods, features for statistical thresholding, common datasets used for document image analysis, experimental results and evaluation and finally, the discussion, future scope and conclusion.

Thresholding Methods

We review thresholding techniques under seven categories: Histogram Shape-Based Methods, Clustering-Based Methods, Entropy-Based Methods, Object Attribute-Based Methods, Spatial Methods, Local Methods and Hybrid Methods.

Histogram Shape-Based Methods

Chen Method: Chen and Wang (2017) propose an adaptive thresholding based on the histogram shape together with one pre- (de-noise) and three post-binarization processes (de-speckle, preserve stroke and improve quality of text region). The histogram's convex hull is calculated and the concavity points which are the deepest points are the potential threshold values. And when the background is similar to the text, logistic sigmoid function is used.

Abdullah Method: In addition to histogram information, this adaptive thresholding algorithm evaluates the behaviour of blobs or objects in an iterative thresholding process. Abdullah *et al.* (2016) claim that the best threshold value lies in the peak value or the highest number of blobs or objects versus every 10 ranges of threshold values starting from 0 until 255. This heuristic thresholding method is tested in a Malaysia license plate recognition system.

Clustering-Based Methods

Yahya method: Yahya *et al.* (2011) propose the following steps in their binarization method: local adaptive equalization and image intensity value process, K-Means clustering to determine the automatic threshold, adaptive thresholding and median filtering. An adaptive threshold value is explored based on a bi-level histogram. The proposed threshold value exploration is carried out by establishing the balance point of the uncertain threshold value, then balancing it with weights closest to the uncertain threshold point until the actual threshold value is found.

Mitianoudis method: Based on the local characteristics of the pixels (i.e., intensity and contrast values), a Local Co-Occurrence Map (LCM) is created from a grayscale document image and is used in the Mixture-of-Gaussian (MoG) clustering model to cluster pixel into two groups: background or character. The Expectation-maximization technique is used to estimate the MoG parameters. This binarization technique is implemented as a 3-stage process: (1) Background stain removal is based on a statistical thresholding of the differences between the text characters and estimated background; (2) Cluster-based binarization and (3) Removal of misclassified components. The proposed binarization method gives comparable performance. It is not the best performing method but it is less complex and fast to run. However, the method is dependent on the

threshold value used in the background stain removal stage. Different document images may require difference threshold value as opposed to a single fixed value (Mitianoudis and Papamarkos, 2015).

Entropy-Based Methods

Li-Lee method: an image thresholding technique is proposed by Li and Lee (1993) whereby the categorization of all the gray levels into two classes is based on the crossed entropy minimization. The means of the two classes, denoted by $\eta_1(t)$ and $\eta_2(t)$, are according to the gray-level t and h_i is the histogram value at the gray-level i . Li and Lee (1993) establish that the optimal threshold T is calculated in order to minimize $\eta(t) = \eta_1(t) + \eta_2(t)$.

Cheng-Chen method: The Cheng-Chen method computes the threshold value based on the maximal entropy principle and the fuzzy C-partition (Cheng *et al.*, 1998). Two fuzzy sets are considered: *Object* and *Background*. Membership functions are then defined for each set. The threshold value is selected as the gray level of which the membership function is equal to 0.5.

Objective Attribute-Based Methods

Pai Method: Pai *et al.* (2010) present a document image binarization algorithm combining both global and local methods for improvement in time complexity and performance. It consists of two steps: block detection and image binarization. An input document image is first segmented into several regions. For each region, the gray-level region is compared to the binarized region. Then, a threshold surface is constructed based on the diversity and the intensity of each region to derive the final binary image.

Tung method: Tung and Wu (2017) propose a binarization technique based on boundary connectivity to produce a threshold surface to binarize gray-level images. First, Sobel operation of a predefined threshold is performed to obtain a binarized image that is used as the global feature evaluation standard of the gray level image. The gray level image is then partitioned into a fixed number of sub-images. The binarization effect is evaluated for each sub-image. The sub-image with the best binarization effect is set as the starting binary area of the gray level image. The binary area is extended by holding and maximizing the boundary connectivity. Neighbouring sub-images that do not satisfy the constraint of boundary connectivity are marked as "ignored" and remain gray. The process is repeated until all sub-images are either binarized or "ignored". Finally, the threshold surface is constructed based on the binarized sub-images of the gray level image.

Ntirogiannis Method: This method also presents a hybrid local adaptive binarization methods at the level of connected component (Ntirogiannis *et al.*, 2014). Their method starts with a background estimation which is

applied together with image normalization for background compensation. This is followed by a global binarization which is performed on the normalized image and noise removal by discarding the very small components. Stroke width and contrast characteristics are then computed as a representative of a document image. Next, local adaptive binarization is performed on the normalized image taking into account the stroke width and contrast characteristics. Lastly, at the connected component level, the two binarized outputs are integrated.

Spatial Methods

Sehad Method: Sehad *et al.* (2013) consider the texture features (e.g. Haralick parameters such as contrast with four distance vectors, etc.) from neighboring pixel information represented as the spatial Gray Level Co-Occurrence Matrices (GLMC) to perform binarization for degraded document images. Image characteristics are estimated based on the joint probability of gray level occurrences of two pixels in a defined direction and distance. Ultimately, a threshold value is computed for each pixel.

Kasmin Method: Kasmin *et al.* (2017) implement a supervised machine learning binarization technique to classify the pixel (i.e. the hotspot pixel) either as foreground/text (1) or background (0) pixel. They extend Sampe *et al.*'s (2011) pixel feature of 4 neighborhood gray level values to an ensemble of steerable local neighborhood gray level information. Eight 3×3 steerable filters based on different orientations: 0° , 45° , 90° , 135° , 180° , 225° , 270° and 315° were used to construct the sets of normalized intensities of gray level values. The probabilities of the pixel being foreground or background from the eight filters are combined using two ensemble techniques: weighted product rule and weighted addition rule respectively to obtain the final pixel value of 1 or 0.

Local Methods

Niblack method: The Niblack algorithm (Niblack, 1985) calculates a local threshold for each pixel by gliding a rectangular window over the whole image. The threshold T is calculated by using the mean m and the standard deviation σ of all pixels in the window.

Sauvola method: The Sauvola algorithm is a modification of that of Niblack (Sauvola *et al.*, 1997), in order to improve binarization performance of documents with a background containing a light texture or varied and uneven illumination.

Wolf method: In order to address Sauvola algorithm problems (low contrast, disparity of gray-level, etc.), Wolf *et al.* (2002) propose a way to normalize the contrast and the image gray-level mean and calculate the threshold adaptively.

Nick method: This method improves the binarization of "white" pages and lighted images and in the case where

the image presents low contrast, by downwards moving of the binarization threshold (Khurshid *et al.*, 2009).

Bataineh Method (a): This proposed two-step method is an adaptive local binarization technique for document images (Bataineh *et al.*, 2011). The two steps are: (1) dynamic partitioning of a document image into binarization windows and (2) application of thresholding method on those windows. This method tackles the binarization challenges such as thin pen stroke lines and low-contrast images.

Bataineh Method (b): This method is based on the variance between pixel contrast (Bataineh *et al.*, 2015). It is a four-stage method: (1) pre-processing, (2) geometrical feature extraction, (3) feature selection and (4) post-processing. In the pre-processing phase, the median filter and histogram equalization are used, then the geometrical features are extracted and feature selection is done for representing the regions. After that, the true text regions are retained and the noise regions are removed. Finally, the post processing is applied for text body reconstruction.

Singh Method: This method (Singh *et al.*, 2012) describes a locally adaptive thresholding technique that removes background by using local mean and mean deviation. Typically, the local mean computational time is highly associated to the window size. This technique also uses integral sum image as a prior processing to calculate local mean.

Shijian Lu Method: In this technique (Lu *et al.*, 2010), a document background surface is firstly estimated through an iterative polynomial smoothing procedure. The text stroke edge is further detected from the compensated document image by using L1-norm image gradient. Finally, the document text is segmented by a local threshold that is estimated based on the detected text stroke edges.

Gatos method: Gatos *et al.* (2008) present an adaptive approach for the binarization. Their method employs pre-processing, a set of binarization methods and post-processing. In the pre-processing phase, they use an adaptive Wiener method based on statistics estimation from a local neighbourhood. In the binarization phase, they combine N binarization results of their previous binarization methodologies (Gatos *et al.*, 2008; 2009). Then, Canny edge detector is used. It inherits from Sobel masks in order to find the edge magnitude of the gray scale image and uses non-Maxima suppression and hysteresis thresholding subsequently. In the post-processing phase, they apply mathematical morphology techniques such as a conditional dilation, shrink and swell filters in sequence.

Lu method: A three-step local binarization method is proposed in Lu *et al.* (2016) which consists of (1) area partition to divide grayscale image into no significant, significant and comparatively significant areas, (2) contrast enhancement to modify the contrast between

foreground and background and lastly (3) local thresholding estimation where the threshold value is calculated as the mean of the highest frequency gray values for the foreground and background. The enhanced contrast improves the local method especially in dealing with document images with thin pen strokes and of low contrast quality.

Hybrid Methods

Kuo Method: Kuo *et al.* (2010) propose a binarization approach based on a hybrid color quantization process, which adaptively takes the global and local image characteristics into account. This method firstly computes the global thresholding value for the entire image. The next step is to make a decision on whether the global thresholding is applicable to the target pixel. If it is inapplicable, then it will apply a local thresholding approach instead.

Moghaddam Method: This method presents an adaptive binarization method inspired by Otsu's method (Moghaddam and Cheriet, 2012). The method, called AdOtsu, uses the Estimated Background (EB) as a priori information to differentiate between text and non-text regions. The estimated background values are calculated in a bootstrap process with implicitly incorporating the binarization method. Also, a priori structural information, including the average stroke width and the average text height, are used to adapt his method with lesser parameter in the document input image. In the post processing corrections, both topological and clustering are used to improve the final binary output.

Bolan Su Method: This method presents a classification framework by combining different thresholding methods (Su *et al.*, 2011). This framework divides the document image pixels into three sets, namely, foreground, background and uncertain pixels. A rule based classifier is then applied iteratively to classify those uncertain pixels into foreground and background, based on the pre-selected foreground and background sets. In the binarization process, the results are then generated by combining existing binarization methods (Otsu, 1975; Sauvola *et al.*, 1997).

Kavallieratou Method: This method presents a hybrid approach that combines global and local thresholding for historical document image (Kavallieratou and Stathis, 2006). Initially, it applies a global thresholding method and then, it identifies the image areas that still contain noise and reapply the global threshold. In short, this technique consists of the following steps: (i) apply Iterative Global Thresholding (IGT) onto document image, (ii) detect the areas with remaining noise and (iii) re-apply IGT for each detected area.

Neves and Mello Method: This algorithm is divided into three phases (Neves and Mello, 2011). In the beginning stage, it identifies the main objects of the

image. The second phase divides the image into sub-images according to the identification phase. The last phase evaluates a local threshold for each sub-image and proceeds with the binarization of each sub-image region. The thresholding steps are written as the following:

- i. Calculate the mean and median value of the sub-image
- ii. Calculate the threshold:

$$T_{nm} = \frac{(\text{mean} + \text{median})}{2}$$

- iii. Transform into binary for each sub-image using T_{nm} threshold value

Pirahansiah Method: Pirahansiah *et al.* (2013) use Peak Signal Noise Ratio (PSNR) element as a consideration to select the best threshold value. PSNR measures the similarities between the original grayscale image and the binarized image – higher PSNR indicates higher similarity. Hence, this method computes the PSNR of the original image to the binarized version at a series of threshold values in increments of 5, within the range of 1 to 256. The corresponding threshold value is selected if it falls between two parameters k_1 and k_2 . These parameters are heuristically chosen at 16 and 10 respectively.

Ismail Method: In the proposed method, Ismail and Abdullah (2014) address problems and weaknesses of the simple methods. They develop a capable method for dealing with the challenging low contrast problem between the foreground and background that causes text information loss. The method consists of image enhancement, proposed local thresholding for binarization and smoothing filter as post-processing for binary image.

Khatatneh Method: A hybrid method is proposed to leveraged on the strengths of global and local methods, aiming to deal with uneven illumination, low contrast, smears, noise and shadows in document images (Al-Khatatneh *et al.*, 2015). Given an input document image, the global standard deviation is calculated. The image is then segmented into sub-images of 40×40 window size and the local standard deviation for each sub-image is computed. The suitable threshold value is determined depending on the global and local standard deviation as shown in Table 1.

While we categorize the techniques, some techniques can fall under more than one category whereby combinations of methods are used especially the newer methods in addressing the degradations present in document images. Table 1 compiles the reviewed techniques together with the respective thresholding functions.

Table 1: Mathematical formulation of thresholding methods

Method and author	Thresholding type	Thresholding function
Cheng-Chen method (Cheng <i>et al.</i> , 1998)	Global/Entropy-based	$U_{object} = \begin{cases} 1, x \leq a \\ (x - c) / (a - c) a < x < c \\ 0, x \geq c \end{cases}$ $U_{background} = \begin{cases} 0, x \leq a \\ (x - a) / (c - a) a < x < c \\ 1, x \geq c \end{cases}$ <p>where x is the independent variable, a and c are parameters determining the shape of the membership functions</p>
Li-Lee method (Li and Lee 1993)	Global/Entropy-based	$N^{1(t)} = \sum_{j=1}^{t-1} i h_j \log\left(\frac{i}{U^1(t)}\right) \text{ And } N_2(t) = \sum_{j=1}^{t-1} i h_j \log\left(\frac{i}{U_2(t)}\right)$ $U_i(t) = \frac{\sum_{i=1}^{t-1} i h_i}{\sum_{i=1}^{t-1} h_i} \text{ et } U_2(t) = \frac{\sum_{i=1}^{256} i h_i}{\sum_{i=1}^{256} h_i}$
Abdullah Method (Abdullah <i>et al.</i> , 2016)	Global/Hsitogram shape-based	$B(x, y) = \begin{cases} 0 & 1(x, y) \leq \Omega_{peak} \\ 1 & otherwise \end{cases}$ $\Omega_{peak(0,1...p-1)} = \begin{cases} peak \text{ if } (\Omega_c > \Omega_{c-1}) \cap (\Omega_c > \Omega_{c+1}) \\ non - peak & otherwise \end{cases}$ <p>where, Ω is the number of objects at $c \times 10$ threshold value, c is the current counter, Ω_{peak} is a series of the peak value derive from Ω_c, p is the total number of array, I is the source image and B is the binary image at location (x, y).</p>
Niblack method (Niblack, 1985)	Local	$T = m + k * \sigma$ <p>where, m is the mean value and σ is the standard deviation value of the pixels inside the window and the value of k is 0.2.</p>
Sauvola method (Sauvola <i>et al.</i> , 1997)	Local	$T = m \cdot \left(1 - k \cdot \left(1 - \frac{\sigma}{R} \right) \right)$ <p>where k is a factor in the range [0.2, 0.5] and R is a gray-level range value</p>
Wolf method Wolf <i>et al.</i> (2002)	Local	$T = (1 - k) \times m + k \times M \times \frac{\sigma}{R} (m - M)$ <p>where m and M is the mean value and σ is the standard deviation value of the pixels inside the window and global image respectively, where k is a factor in the range and R is a gray-level range value</p>
Gatos method (Gatos <i>et al.</i> , 2008)	Local	$B_R(x, y) = \begin{cases} 1, \text{if } \sum_{i=1}^{2m+1} R_i(x, y) > m \\ 0, otherwise \end{cases}$ <p>where, R is represents the result of m a set of binarization methods</p>
Nick method Khurshid <i>et al.</i> , 2009)	Local	$T = m + k \sqrt{\frac{(\sum p_i^2 - m^2)}{NP}}$ <p>where k is a factor in the range [-0.1,-0.2], p_i is the pixel value of the gray-scale levels in the image, and NP is the total number of pixels in the image:</p>
Ismail method (Ismail and Abdullah, 2014)	Local	$M_L = \frac{I(x, y)}{W_{size}}$ $TH_W = M_L \left(\frac{(1 - T)}{K} \right)$ <p>where M_L is mean value for all pixels, $I(x, y)$ is the gray image and W_{size} is the window size and TH_W is the local threshold value,</p>

Table 1: Continue

		is the local mean value for $I(x, y)$ image pixels based on W size, k is a constant value (we suggest $k = 100$ for obtaining the best result)
Bataineh method (Bataineh <i>et al.</i> , 2015)	Local	$Th = \frac{m^2 * \sigma}{(m + \sigma)(127 + \sigma)}$ <p>where T is the thresholding value, m is the mean value of the global image pixels, r is the standard deviation of the image pixels and 127 is the median of the gray-level value (255)</p>
Lu method (Lu <i>et al.</i> , 2016)	Local	$T = \frac{ff_{foreground} + ff_{background}}{2}$ <p>where $ff_{foreground}$ and $ff_{background}$ refer to the highest frequency gray value in the foreground and background respectively</p>
Schad method (Schad <i>et al.</i> , 2013)	Local/Spatial (pixel-wise adaptive)	$T(i, j) = m(i, j) + k * a$ <p>where $a = contract(GLMC) = \sum_{i=0}^n \sum_{j=0}^n (i - j) * GLMC(i, j)$ a is a parameter based on Haralick descriptor and $GLMC(i, j)$ are elements of the $GLMC$ matrix at i and j coordinates and n is the number of image gray levels, m is the local mean of image pixels within a centered window and k is a parameter defined by experiment</p>
Chen Method (Chen <i>et al.</i> , 2017)	Local / Histogram shape-based	$d(I'(x, y)) = q\delta \left[\frac{(1 - p_2)}{1 + \exp\left(\frac{-4I'(x, y)}{b(1 - p_1)} + \frac{2(1 + p_1)}{1 - p_1}\right)} + p_2 \right]$ <p>where d is the threshold for pre-processed image I' using the logistic sigmoid function for when the background is similar to the text. q is the weighting parameter (set at 0.6 by experiment), δ is the average difference between black and white pixels, p_1 and p_2 are</p>
Singh method (Singh <i>et al.</i> , 2012)	Hybrid	$T(x, y) = m(x, y) \left[1 + k \left(\frac{\partial(x, y)}{1 - \partial(x, y)} - 1 \right) \right]$ <p>where $\partial(x, y) = I(x, y) - m(x, y)$ is the local mean deviation and k is a bias which can control the level of adaptation varying threshold value within $[0, 1]$ only</p>
Shijian Lu method Lu <i>et al.</i> (2010)	Hybrid	$R(x, y) = \begin{cases} 0 & \text{if } (N_e \geq N_{min} \cap \& \& I(x, y) \leq E_{mean}) \\ 1 & \text{otherwise} \end{cases}$ $E_{mean} = \frac{\sum_{neighbor} I(x, y) \times I - E(x, y)}{N_e}$ <p>where $I(x, y)$ refers to the normalized document image, N_e refers to the number of the detected stroke edge pixels within a local neighborhood window, N_{min} denotes a threshold that specifies the minimum number of detected stroke edge pixels (within the neighborhood window) of possible text pixel, and E_{mean} refers to the mean image intensity of the detected stroke edge pixels within the local neighborhood window</p>
Kuo method (Kuo <i>et al.</i> , 2010)	Hybrid	$Th_{local} = \begin{cases} \mu_{local} \frac{\sigma}{Th_{Global}}, g(p) - \mu_{local} \geq \frac{Th_{Global}}{16} \\ \mu_{local} \times \left[1 + k \left(\frac{\sigma}{Th_{Global}} - 1 \right) \right], g(p) - \mu_{local} \geq \frac{Th_{Global}}{16} \end{cases}$ <p>where μ is local mean, σ is local standard deviation, k is a free parameter and $g(p)$ is a gray pixel</p>

Table 1: Continue

Bolan Su method (Su <i>et al.</i> , 2011)	Hybrid	$P(x) = \begin{cases} foreground, \frac{con(x)}{con_F} > \frac{con_B}{con(x)} \parallel \frac{I_F}{I(x)} > \frac{I(x)}{I_B} \\ background, otherwise \end{cases}$ <p>where $P(x)$ is uncertain pixel, $con(x)$ is contrast and $I(x)$ is intensity of feature, respectively, con_F and con_B are the mean contrast and I_F and I_B are intensity feature value of foreground and background pixels inside a local 3×3 neighbourhood window, respectively</p> $FB(x, y) = CO(x, y) \cup (O(x, y).f(x, y))$
Ntirogiannis method (Ntirogiannis <i>et al.</i> , 2014)	Hybrid/ Objective attribute-based	$f(x, y) = \begin{cases} 1 & \text{if } \sum_{m=x-1}^{x+1} \sum_{n=y-1}^{y+1} CO(m, n) > 0 \\ 0 & \text{otherwise} \end{cases}$ $CO(x, y) = \bigcup_{j=1}^n NB_j(x, y), \quad \forall j : d(j) = true$ <p>where $FB(x, y)$ is the final binarization, $O(x, y)$ is connected component of Otsu, a pixel of the combination result image $CO(x, y)$ can be detected in the 8-connected neighbourhood, and NB is the connected component of the Niblack result</p>
Kavallieratou method (Kavallieratou and Stathis, 2006)	Hybrid	$f(s) > m + k \times \sigma$ <p>where $f(s)$ is the frequency of segment S, $k = 2$ is a free parameter, m is local mean, σ is the standard deviation of the black pixel frequency in the image</p>
Bataineh method (Bataineh <i>et al.</i> , 2011)	Hybrid	$T_w = mw - \frac{m_g^2}{(m_g + \sigma_w)(\sigma_{adaptive} + \sigma_w)}$ <p>where T_{con} is the confusion threshold of the whole document image, m_g is the global mean value of the image pixels; \max_{level} is the max level value of the image and σ_g is the standard deviation of the image pixels</p> <p>if $G > L$, then</p>
Khatatneh method (Al-Khatatneh <i>et al.</i> , 2015)	Hybrid	$T = m * \left(1 - k * \left(1 - \frac{Abs(L - G)}{128} \right) \right)$ <p>Else</p> $T = m - f * sqrt(L^3 + m^2)$ <p>Depending on the global standard deviation (G) and the local standard deviation (L) of each window, the threshold value, T is calculated where $k = 0.5$ and $f = 0.7$ are the parameters used for determining the number of edge pixels that are considered as object pixels. k and f values are the values that produce the best results and are heuristically obtained</p>
Tung method (Tung and Wu, 2017)	Hybrid/Objective attribute-based	$TH_{(x,y)} = \sum_{t=1}^k \frac{1}{N_t \cdot DIS} * N_t \cdot TH$ <p>where $TH(x, y)$ is the threshold value of (x, y) in the threshold surface, N_t, TH and $N_t \cdot DIS$ the threshold and the distance of for (x, y) respectively</p>

Statistical Thresholding Features

A major concern in object detection and recognition is to select features that are able to describe the object of interest (which is the text in document images in this context). As can be summarized from the literature, most binarization

algorithms require global and local information for image description. Quite a number of them used histogram projection and statistical parameters (first, second or multi orders) for image description (Kuo *et al.*, 2010; Li and Lee 1993; Otsu, 1975). Table 2 summarizes the popular features/parameters used in the state-of-the-art thresholding techniques.

Table 2: The most global and local statistical information description features of image used in literature thresholding techniques

Thresholding information description	Type feature	Description
Geometric properties	Size position	Summation of all pixel values or the number of pixels of the object The center area of the object. In order to calculate the position of the object, we need to calculate the first order moments
	Orientation	Widely employed in the edge and line detection, which is an important step in the object detection
	Histogram projection	Splitting the range of the gray level into equal-sized classes. For each class, the number of pixels from the image that fall into each class is counted
Basic statistical distribution features	Mean and weighted average	Obtained by dividing the sum of observed values by the number of observations
	Median	It is calculated by listing the pixel values in an ascending order, then finding the point that is exactly in the middle
	Mode	The number of pixel values with the highest frequency
	Standard deviation and Weighted standard deviation	Close entire set of data to the average value (small standard deviation tightly grouped) whereas weighted standard deviation normally uses a constant value to multiply with local or global standard deviation value
	Standard deviation of the Mean	Compute the standard deviation and mean for each sub image
	Variance	The variance is defined as the average value of the quantity (The square of distance from mean)
	Moments	The moment in general used to weighted average of the image pixels' intensities
Entropy-Based method	Cross-Entropy (CE) Method	Cross entropy measures the information theoretic distance between two distributions
	Maximum entropy model	An automatic thresholding method where the optimal threshold value can be found by maximizing the entropy of the resulting classes (foreground and background)
Texture features	Edge detection	Significant local changes of intensity in an image. Edges typically occur on the boundary between two different regions. This is to find the object boundary
	Co-occurrence matrices	Captures numerical features of a texture using spatial relations of similar gray region
	Blob distribution	The number of blobs or objects in the iterative thresholding and select the highest number blob at a particular threshold as the best threshold value

Datasets

There are several common datasets used for document image analysis. The most important datasets are as follows.

DIBCO

DIBCO which stands for Document Image Binarization Contest (DIBCO, 2009; 2010; 2011; 2012; 2013) is organized by the International Conference on Document Analysis and Recognition (ICDAR). Datasets used in the series of DIBCO contests are available for download at: <https://vc.ee.duth.gr/dibco2017/>.

These DIBCO datasets consist of different types of document images ranging from scanned machined-printed images to scanned handwritten images. The datasets ensure wide coverage of the various types of challenges to the process of binarization (degradations). Both color and gray-scale images are inclusive in the datasets.

H-DIBCO 2014 and 2016 datasets which consist of handwritten images only are also available in the given link above.

The Degraded Document Images

This dataset presents a selection of degraded documents images representing the Arabic calligraphy and Latin scripts. It contains 18 document images: 14 Arabic calligraphy and 4 Latin scripts. Each document image presents a set of binarization challenges such as dirty spots, low contrast between text and background, seeping ink between documents and multi-colored text. This dataset was collected from books and historical manuscripts (Bataineh *et al.*, 2012) and has been used to evaluate the performance of various binarization methods.

Document Binarization Dataset (DBD)

This complete dataset is composed of document images, ground truth and tools to perform an evaluation of

binarization algorithms. It allows pixel-based accuracy and OCR-based evaluations. It is available for download at: <https://www.lrde.epita.fr/wiki/Olena/DatasetDBD>.

There are three types of document images of varying quality in this dataset:

- Original: To evaluate the binarization quality on perfect documents mixing text and images.
- Clean: To evaluate the binarization quality on a perfect document with text only.
- Scanned: To evaluate the binarization quality on slightly degraded documents with text only.

Experimental Result and Evaluation

Binarization performance is typically evaluated by conducting experiments. In this section, we present the performance results reported by the above aforementioned methods in Table 3. Most of the

methods (Al-Khatatneh *et al.*, 2015; Bataineh *et al.*, 2011; 2015; Chen and Wang, 2017; Gatos *et al.*, 2009; Lu *et al.*, 2010; 2016; Moghaddam and Cheriet, 2012; Neves and Mello, 2011; Pratikakis *et al.*, 2013; Su *et al.*, 2011) are evaluated based on one or more of the benchmark metrics adopted by DIBCO 2009. These metrics include the F-Measure, Picture Signal-to-Noise Ratio (PSNR), Negative Rate Metric (NRM) and Misclassification Penalty Metric (MPM). Interested readers may refer to Gatos *et al.* (2009) for the full description of these metrics.

Some researchers use an ensemble of measurements – in addition to the four metrics mentioned, other metrics used include Mean Square Error (MSE), computational time, average quality score, p-F-measure (p-FM) and Distance Reciprocal Distortion (DRD) (Mitianoudia and Papamarkos, 2015; Ntirogiannis *et al.*, 2014; Pai *et al.*, 2010; Singh *et al.*, 2012) to indicate the robustness of their approaches.

Table 3: The results for 22 methods in document image binarization

Method and author	Result	Method and author	Result
Cheng-Chen method (Cheng <i>et al.</i> , 1998)	Very well for preserving the main features of the components images.	Singh method (Singh <i>et al.</i> , 2012)	Visual evaluation and Computational Time is measured to be ~0.234 Sec. Achieved lowest computational time compared Niblack, Sauvola and Bernsen methods
Li-Lee method (Li and Lee, 1993)	average quality score result = 0.114	Kuo method (Kuo <i>et al.</i> , 2010)	Competitive visual result compare to Niblack, Sauvola, Chang and Otsu methods
Niblack method (Niblack, 1985)	Distance between resulting binarization and Ground truth images = 146	Bolan Su method (Su <i>et al.</i> , 2011)	Using the F-Measure method: The Combined results of Otsu's and Sauvola's =86.62%. Lu's method and Su's = 93.18%
Sauvola method (Sauvola <i>et al.</i> , 1997)	Distance between resulting binarization and Ground truth images 28.7	Shijian Lu method (Lu <i>et al.</i> , 2010)	Using F - measure method = 91.24 %
Wolf method (Wolf <i>et al.</i> , 2002)	Distance between resulting binarization and Ground truth images 53.16	Ntirogiannis method (Ntirogiannis <i>et al.</i> , 2014)	F-measure, PSNR, NRM, MPM, p-F- measure (p-FM) and Distance Reciprocal Distortion (DRD) Top performance in most cases
Nick method (Khurshid <i>et al.</i> , 2009)	Characters correctly recognized = 99.41	Kavallieratou method (Kavallieratou and Stathis, 2006)	Precision is calculated based on the 43 images already in good condition = 97.67%.
Neves and Mello Method (Neves and Mello, 2011)	F-measure, PSNR, NRM, MPM, GA respectively Result = 88.7052, 18.7090, 0.0576, 0.6823, 0.9360	Moghaddam Method (Moghaddam and Cheriet, 2012)	Using the F - measure method automatic mode = 91.57%. Grid-based AdOtsu = 92.01%. Multiscale grid-based AdOtsu = 92.06%
Gatos method (Gatos <i>et al.</i> , 2009)	F - Measure method result = 91.9 %.	Pai Method (Pai <i>et al.</i> , 2010)	Average processing time = 0.351. Average Recognition Rate = 98.22%
Bataineh method (a) (Bataineh <i>et al.</i> , 2011)	F-mean, PSNR, NRM respectively = 10.5, 6.16, 89.34	Bataineh method (b) (Bataineh <i>et al.</i> , 2015)	F-mean, PSNR, NRM respectively = 17.5 5.14, 88%
Mitianoudis method (Mitianoudia and Papamarkos, 2015)	PSNR, Mean-Square Error (MSE), Recall, Precision, F-mean, NRM Without contrast, results respectively = 15.29, 0.0295, 0.8331, 0.9466, 0.8862, 0.0872	Chen method (Chen <i>et al.</i> , 2017)	PSNR, F-mean respectively = 18.2381, 95.7 Experimentally shown to outperform Niblack, Sauvola, Gatos and 3 other methods not reviewed in this article
Lu method (Lu <i>et al.</i> , 2016)	F-mean, PSNR and NRM for handwritten images are 82.82% 11.86 and 8.76 respectively; for printed text images are 87.12%, 10.44 and 8.92 respectively Experimentally shown to outperform Niblack, Sauvola and Nick methods in term of F-measure and PSNR for both handwritten and printed text	Khatatneh method (Al-Khatatneh <i>et al.</i> , 2015),	F-mean and PSNR, for handwritten images are 79.63% and 16.56 respectively; for printed images are 87.6% and 15.94 respectively Experimentally shown to outperform Niblack, Otsu, Sauvola, Nick and Wolf methods

Other methods (Khushid *et al.*, 2009; Niblack, 1985; Sauvola *et al.*, 1997; Wolf *et al.*, 2002) evaluate the accuracy of their works using Optical Character Recognition (OCR) accuracy testers which are limited to character sets supported by OCR systems and methods (Kuo *et al.*, 2010; Singh *et al.*, 2012) simply perform a visual evaluation which could be bias. Also, when comparisons of methods have been performed by researchers, it is indicated in the result column.

Discussion

Most of the binarization techniques in image analysis are based on statistical parameters. Such methods extract statistical values based on the spatial distributions of gray level values in the image. Statistical methods are classified based on statistical orders. The first-order statistics find the value of each level individually and extract the properties based on those values. The second-order statistics find the value of two levels by relating the first-order values with some geometrical relationship and indicate these levels as important features. The third- and higher-order statistics depend on finding values of compound properties of the image.

The first-order statistics are based on individual levels. The same image may consist of different brightness based on different cases. These different cases lead to different features for the same image. Therefore, the first-order statistic is affected by the scanning machine and its settings, in addition to the distortion and noises in image. The higher-order statistics may need several methods to find several properties; therefore, requiring more features leading to higher complexity in the algorithm.

While global methods are simple and easy to implement, these methods are not capable of handling poor quality document images (e.g. high noise, poor contrast, low illuminations, spots and patches, etc) – the images' characteristics are not leveraged. On the other hand, in local thresholding methods, the threshold values are determined locally either pixel by pixel (Kasmin *et al.*, 2017; Niblack, 1985; Sehad *et al.*, 2013) or region by region (Neves and Mello, 2011; Pai *et al.*, 2010; Tung and Wu, 2017). Then, a specified region can have a threshold value that is changed from region to region according to the threshold candidate selection for a given area. The main advantage of local methods is that in degraded document images where considerable background noise or variation in contrast and illumination is present; there exists many pixels that cannot be easily classified as foreground or background by global methods. In such cases, binarization with local thresholding is more appropriate and accurate.

These local methods are typically concerned with window size and local features (e.g. contrast, variance, position, orientation, edge, connectivity, texture etc.). Small binarization windows are effective in removing

noise but may break up large text whereas large binarization windows are more effective in preserving the text but do not remove noise well. A fixed set of parameter values by the user is also impractical since each image has different levels of noise. Hence, the window size and parameter value determination need to be dynamic according to the characteristics of the input image to strike a balance between noise removal and text preservation (Bataineh *et al.*, 2011; Boiangiu *et al.*, 2011; Lazzara and Géraud, 2014). Such binarization methods result in algorithms that are time consuming with high computational cost.

The recent trend of hybrid methods attempt to reduce the complexity by leveraging on the strengths of global and local thresholding, that is, improved adaptability of handling various kinds of noise at different areas but with lower complexity. The reviewed hybrid methods are generally combinations and enhancements of existing global and local methods (Kuo *et al.*, 2010; Moghaddam and Cheriet, 2012; Su *et al.*, 2011) as well as novel hybrid approach such as Khatatneh method that tackles uneven illumination, shadows, low contrast, smears and heavy noise (Al-Khatatneh *et al.*, 2015).

Future Scope

Extracting text from document images especially of degraded quality remains a challenging problem – no one method can address all types of degradations present in document images. The hybrid adaptive method may be the solution that is able to solve some of the more complex binarization challenges where it is implemented with successive post-binarization enhancement processes (e.g. de-speckle, stroke preservation, Gaussian and median filters etc.) to improve the binarization results (Al-Khatatneh *et al.*, 2015; Chen and Wang, 2017; Nafchi *et al.*, 2014).

Conclusion

In this study, we have reviewed several classes of document image binarization methods that have been proposed in the literature, mainly to transform gray-level or colored document images into binary document images. We have presented some of documents images thresholding techniques that are most frequently cited in the literature and are the basis for most of the modern binarization approaches. Some techniques deal with certain degradation better than other techniques. Clearly, there is still not one method that can address all the degradations found in existing document images which range from handwritten to typed written document images to ancient, historical to modern document images. We have explained and tabulated the thresholding equations and statistical parameters used in the reviewed methods to identify suitable thresholding value(s). Furthermore, this

article provides information describing available datasets and highlighting state-of-the-art evaluation measures and metrics. There are three different approaches to evaluating the binarization techniques: visual evaluation, OCR approach and metrics adopted by DIBCO which is most commonly adopted by most researchers. This survey serves as a guide to researchers in the area for future advancement in statistical binarization techniques for degraded document images.

Acknowledgment

The authors would like to thank the Faculty of Information Science and Technology, Center for Artificial Intelligence Research and Center for Cyber Security of the Universiti Kebangsaan Malaysia for providing the facilities and financial support under the DIP-2015-023 grant entitled "Object Descriptor via Optimized Unsupervised Learning Approaches", PRGS/1/2016/ICT02/UKM/02/1 grant entitled "Intelligent Vehicle Identity Recognition for Surveillance" and AP-2017-005/2 grant entitled "Self Crime Prevention".

Author's Contributions

Saad M. Ismail: Performed the review of the binarization techniques, drafted and revised the article.

Siti Norul Huda Sheikh Abdullah: Supervised the work, provided critical feedback and revised the article.

Fariza Fauzi: Performed the critical revision of the article and reviewed the latest binarization techniques.

Ethics

There are no ethical issues or conflict of interest.

References

- AL-Khatatneh, A.M., S.A. Pitchay and M.K. Al-qudah, 2015. Compound binarization for degraded document images. *ARPN J. Eng. Applied Sci.*, 6608: 594-599.
- Bataineh, B., S.N.H.S. Abdullah and K. Omar, 2011. An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows. *Pattern Recognition Lett.*, 32: 1805-1813.
- Bataineh, B., S.N.H.S. Abdullah and K. Omar, 2012. A novel statistical feature extraction method for textual images: Optical font recognition. *Expert Syst. Applic.*, 39: 5470-5477.
- Bataineh, B., S.H.S. Abdullah and K. Omar, 2015. Adaptive binarization method for degraded document images based on surface contrast variation. *Pattern Anal. Applic.*
- Boiangiu, C.A., A. Olteanu, A. Stefanescu, D. Rosner and N. Tapus *et al.*, 2011. Local thresholding algorithm based on variable window size statistics. *Proceedings of the 18th International Conference on Control Systems and Computer Science (SCS' 11)*, 2: 647-652.
- Chen, Y. and L. Wang, 2017. Broken and degraded document images binarization. *Neurocomputing*, 237: 272-280.
- Cheng, H.D., J.R. Chen and J. Li, 1998. Threshold selection based on fuzzy c-partition entropy approach. *Pattern Recognition*, 31: 857-870.
- Gatos, B., K. Ntirogiannis and I. Pratikakis, 2009. ICDAR 2009 Document Image Binarization Contest (DIBCO 2009). Paper read at ICDAR.
- Gatos, B., I. Pratikakis and S.J. Perantonis, 2008. Efficient binarization of historical and degraded document images. *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems*, Sept. 16-19, IEEE Xplore Press, Nara. DOI: 10.1109/DAS.2008.66
- Ismail, S.M. and S.N.H.S. Abdullah, 2014. Novel binarization method for enhancing ancient and historical manuscript images. *Proceedings of the Mexican International Conference on Artificial Intelligence*, Springer International Publishing.
- Kasmin, F., A. Abdullah and A.S. Prabuwo, 2017. Ensemble of steerable local neighbourhood grey-level information for binarization. *Pattern Recognition Lett.* 98: 8-15.
- Kavallieratou, E. and S. Stathis, 2006. Adaptive binarization of historical document images. *Proceedings of the 18th International Conference on Pattern Recognition (ICPR' 06)*.
- Kefali, A., T. Sari and M. Sellami, 2010. Evaluation of several binarization techniques for old Arabic documents images. *Proceedings of the International Symposium on Modeling and Implementing Complex Systems, (ICS' 10)*.
- Khurshid, K., I. Siddiqi, C. Faure and N. Vincent, 2009. Comparison of niblack inspired binarization methods for ancient documents. *IS&T/SPIE Electron. Imaging*.
- Kuo, T.Y., Y.Y. Lai and Y.C. Lo, 2010. A novel image binarization method using hybrid thresholding. *Proceedings of the IEEE International Conference on Multimedia and Expo (IME' 10)*.
- Lazzara, G. and T. Géraud, 2014. Efficient multiscale sauvola's binarization. *Int. J. Document Anal. Recognition*, 17: 105-123.
- Li, C.H. and C.K. Lee, 1993. Minimum cross entropy thresholding. *Pattern Recognition*, 26: 617-625.
- Lu, D., X. Huang, C. Liu, X. Lin and H. Zhang *et al.*, 2016. Binarization of degraded document image based on contrast enhancement. *Proceedings of the 35th Chinese Control Conference*, Jul. 27-29, IEEE Xplore Press, Chengdu, pp: 4894-4899. DOI: 10.1109/ChiCC.2016.7554113

- Lu, S., B. Su and C.L. Tan, 2010. Document image binarization using background estimation and stroke edges. *Int. J. Document Anal. Recognition*, 13: 303-314. DOI: 10.1007/s10032-010-0130-8
- Mahajan, P.A. and S.A. Patil, 2015. A review of image binarization techniques for ancient degraded documents. *Int. J. Sci. Res.*, 4: 54-56.
- Mitianoudis, N. and N. Papamarkos, 2015. Document image binarization using local features and Gaussian mixture modeling. *Image Vision Comput.*, 38: 33-51.
- Moghaddam, R.F. and M. Cheriet, 2012. AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization. *Pattern Recognition*, 45: 2419-2431.
- Morse, S.B., 1998. Lecture 2: Image processing review, neighbors, connected components and distance. Brigham Young University.
- Nafchi, H.Z., R.F. Moghaddam M. Cheriet, 2014. Phase-based binarization of ancient document images: Model and applications. *Image Proc. IEEE Trans.*, 23: 2916-2930.
- Neves, R.F.P. and C.A.B. Mello, 2011. A local thresholding algorithm for images of handwritten historical documents. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Oct. 9-12, IEEE Xplore Press, Anchorage. DOI: 10.1109/ICSMC.2011.6084150
- Niblack, W., 1985. *An Introduction to Digital Image Processing*. 2nd Edn., Prentice-Hall, Englewood Cliffs, ISBN-10: 0134806743, pp: 215.
- Ntirogiannis, K., B. Gatos and I. Pratikakis, 2014. A combined approach for the binarization of handwritten document images. *Pattern Recognition Lett.*, 35: 3-15.
- Otsu, N., 1975. A threshold selection method from gray-level histograms. *Automatica*, 11: 23-27.
- Pai, Y.T., Y.F. Chang and S.J. Ruan, 2010. Adaptive thresholding algorithm: Efficient computation technique based on intelligent block detection for degraded document images. *Pattern Recognition*, 43: 3177-3187.
- Pirahansiah, F., H.S. Abdullah and S. Sahran, 2013. Peak signal-to-noise ratio based on threshold method for image segmentation. *J. Theoretical Applied Inf. Technol.*, 57: 158-168.
- Pratikakis, I., B. Gatos and K. Ntirogiannis, 2013. ICDAR 2013 document image binarization contest (DIBCO 2013). *Proceedings of the 12th International Conference on Document Analysis and Recognition*, (DAR' 13).
- Sampe, I.E., H.W. Dann, Y.S. Tsai and C.C. Lin, 2011. Segmentation of mitochondria in fluorescence micrographs by SVM. *Proceedings of the 4th International Conference on Biomedical Engineering and Informatics (BMEI)*, pp: 491-495.
- Sauvola, J., T. Seppanen, S. Haapakoski and M. Pietikainen, 1997. Adaptive document binarization. *Proceedings of the 4th International Conference on Document Analysis and Recognition*, (DAR' 97).
- Sehad, A., Y. Chibani, M. Cheriet and Y. Yaddaden, 2013. Ancient degraded document image binarization based on texture features. *Proceedings of the 8th International Symposium on Image and Signal Processing and Analysis*, Sept. 4-6, IEEE Xplore Press, Trieste. DOI: 10.1109/ISPA.2013.6703737
- Sezgin, M. and B. Sankur, 2004. Survey over image thresholding techniques and quantitative performance evaluation. *J. Electron. Imaging*, 13: 146-168.
- Singh, T.R., S. Roy, O.I. Singh, T. Sinam and K. Singh, 2012. A new local adaptive thresholding technique in binarization.
- Abdullah, S., S.N. Huda, K. Omar, A.S. Zaini and M. Petrou *et al.*, 2016. Determining adaptive thresholds for image segmentation for a license plate recognition system. *Int. J. Adv. Comput. Sci. Applic.*, 7: 1-14.
- Su, B., S. Lu and C.L. Tan, 2011. Combination of document image binarization techniques. *Proceedings of the International Conference on Document Analysis and Recognition*, (DAR' 11).
- Tung, C.H. and Z.L. Wu, 2017. Binarization of uneven-lighting image by maximizing boundary connectivity. *J. Statistics Management Syst.*, 20: 175-196.
- Wolf, C., J.M. Jolion and F. Chassaing, 2002. Extraction de texte dans des vidéos: Le cas de la binarisation. *Proc. RFIA*, 1: 145-152.
- Yahya, S.R., S.N.H.S. Abdullah, K. Omar and C.Y. Liong, 2011. Adaptive binarization method for enhancing ancient malay manuscript images. *Proceedings of the Australasian Joint Conference on Artificial Intelligence*, (CAI' 11), Springer Berlin Heidelberg, pp: 619-627.